"To Meet, Or Not To Meet, That Is The Question"
Optimizing Interaction Strategies

Problem Statement

Company: MFS Investments is an Asset Management Company. In the Retail segment, investors entrust money to financial advisors who, in turn, come to MFS to buy financial products.

Problem: The MFS client-facing team interacts with financial advisors over time to foster a long-term relationship and to sell MFS products. Our goal is to optimize these sequences of interactions (i.e.: interaction strategies)

Approach

- 3-Step Approach: Prepare the Data, Predict and Optimize
- End-to-end Pipeline: From raw data to a final recommendation
- Modular Solution: Blocks can be swapped and implemented independently

Impact

We compare hypothetical outcomes using our approach under different scenarios (i.e.: different assumptions for our optimization model) to current strategies for the main client firm of MFS.

We observe a sizeable expected increase in sales in all scenarios

Phase 1: Prepare the Data

We aggregate sales and interactions by advisor for each quarter

1.5 Million records × 68 Features

We use pretrained embedding and cater them to our corpus to impute VADER scores for unsupervised sentiment analysis

Train Word2Vec
Impute VADER scores
Normalize sentiment by wholesaler

We interpret perform clustering to uncover similarities across clients

Phase 2: Predict Interaction Effects

Models
- Random Forest (Selected)
- KNN Regression
- Decision Trees
- XGBoost
- Lasso

Main Features
- Historical Sales
- Portfolio Returns
- Interaction Sentiment
- Interaction Type
- Interaction Content (to date)

Out-of-sample performance on the subset of clients for whom we optimize ~ 1.2 times the performance of the baseline

Phase 3: Optimize Strategies

We optimize interaction sequences for 6 months over the top client firm. We select the strategies that most frequently occur and validate them with domain experts

Optimization Formulation

\[
\max \sum_{i,j} \left( \sum_{c} \alpha_{ijc} \right) s_{ij} - \lambda \sum_{i,j} m_j s_{ij} - \mu \sum_{i,j} c_j s_{ij}
\]

\[
s.t. \quad \sum_{i,j} s_{ij} \leq 1 \quad \forall i \in [n]
\]

\[
\sum_{i,j} m_j s_{ij} \leq 25 \quad \forall i \in [k]
\]

\[
\sum_{i,j} c_j s_{ij} \leq 50 \quad \forall i \in [k]
\]

\[
s_{ij} \in \{0,1\}
\]

m_j, \lambda, \mu

Cost of meetings and calls

38% Fewer advisors met / called

250 + Million $ Increase in sales

29% Less interactions

We perform interpretable clustering to uncover similarities across clients

Other Client
Company
Yes
No

Sales Frequency

Cluster 0
Cluster 1

Phase 3: Optimize Strategies

We optimize interaction sequences for 6 months over the top client firm. We select the strategies that most frequently occur and validate them with domain experts

Optimization Formulation

\[
\max \sum_{i,j} \left( \sum_{c} \alpha_{ijc} \right) s_{ij} - \lambda \sum_{i,j} m_j s_{ij} - \mu \sum_{i,j} c_j s_{ij}
\]

\[
s.t. \quad \sum_{i,j} s_{ij} \leq 1 \quad \forall i \in [n]
\]

\[
\sum_{i,j} m_j s_{ij} \leq 25 \quad \forall i \in [k]
\]

\[
\sum_{i,j} c_j s_{ij} \leq 50 \quad \forall i \in [k]
\]

\[
s_{ij} \in \{0,1\}
\]

m_j, \lambda, \mu

Cost of meetings and calls

Interpretable Cross validation algorithm

Our algorithm consistently converges to models with better out of sample performance

\[
\text{Performance Convergence}
\]

Phase 1: Prepare the Data

We aggregate sales and interactions by advisor for each quarter

1.5 Million records × 68 Features

We use pretrained embedding and cater them to our corpus to impute VADER scores for unsupervised sentiment analysis

Train Word2Vec
Impute VADER scores
Normalize sentiment by wholesaler

We interpret perform clustering to uncover similarities across clients

Phase 2: Predict Interaction Effects

Models
- Random Forest (Selected)
- KNN Regression
- Decision Trees
- XGBoost
- Lasso

Main Features
- Historical Sales
- Portfolio Returns
- Interaction Sentiment
- Interaction Type
- Interaction Content (to date)

Out-of-sample performance on the subset of clients for whom we optimize ~ 1.2 times the performance of the baseline

Phase 3: Optimize Strategies

We optimize interaction sequences for 6 months over the top client firm. We select the strategies that most frequently occur and validate them with domain experts

Optimization Formulation

\[
\max \sum_{i,j} \left( \sum_{c} \alpha_{ijc} \right) s_{ij} - \lambda \sum_{i,j} m_j s_{ij} - \mu \sum_{i,j} c_j s_{ij}
\]

\[
s.t. \quad \sum_{i,j} s_{ij} \leq 1 \quad \forall i \in [n]
\]

\[
\sum_{i,j} m_j s_{ij} \leq 25 \quad \forall i \in [k]
\]

\[
\sum_{i,j} c_j s_{ij} \leq 50 \quad \forall i \in [k]
\]

\[
s_{ij} \in \{0,1\}
\]

m_j, \lambda, \mu

Cost of meetings and calls

Interpretable Cross validation algorithm

Our algorithm consistently converges to models with better out of sample performance

\[
\text{Performance Convergence}
\]